METHOD FOR DETECTING MISALIGNED PHONETIC UNITS FOR A CONCATENATIVE TEXT-TO-SPEECH VOICE

ABSTRACT

A method of filtering phonetic units to be used within a concatenative text-to-speech (CTTS) voice. Initially, a normality threshold can be established. At least one phonetic unit that has been automatically extracted from a speech corpus in order to construct the CTTS voice can be received. An abnormality index can be calculated for the phonetic unit. Then, the abnormality index can be compared to the established normality threshold. If the abnormality index exceeds the normality threshold, the phonetic unit can be marked as a suspect phonetic unit. If the abnormality index does not exceed the normality threshold, the phonetic unit can be marked as a verified phonetic unit. The concatenative text-to-speech voice can be built using the verified phonetic units.